



---

# Working Life and Human Capital Investment: Causal Evidence from Pension Reform

---

**Niklas Gohl** (DIW Berlin and Potsdam University)

**Peter Haan** (DIW Berlin and FU Berlin)

**Elisabeth Kurz**

**Felix Weinhardt** (DIW Berlin, CESifo, IZA, CEP/LSE)

Discussion Paper No. 212

December 11, 2019

# Working Life and Human Capital Investment: Causal Evidence from Pension Reform\*

Niklas Gohl<sup>†</sup>, Peter Haan<sup>‡</sup>, Elisabeth Kurz and Felix Weinhardt<sup>§</sup>

December 11, 2019

## Abstract

In this paper we present a life-cycle model with human capital investment during working life through training and provide a novel empirical test of human capital theory. Using a sizable pension reform which shifts the retirement age between two adjacent cohorts by three years, we document causal evidence that an increase in the working life increases investment into human capital through training. We estimate this effect using a regression discontinuity design based on a large sample from the German microcensus. We discuss and test further predictions regarding the relation between initial schooling, training, and the reform effect and show that only individuals with a college degree increase human capital investment. Our results speak to a large class of human capital models as well as policies extending or shortening working life.

**Keywords:** human capital, retirement policies, RDD.

**JEL classification:** J24; J26; H21.

---

\*We thank Pio Baake, Rainald Borck, and participants at seminars at the Universities of Cologne, Vienna, and the Lund Workshop on the Economics of Education for comments. We gratefully acknowledge funding from the German Science Foundation through the CRC/TRR190 (Project number 280092119) and Project HA5526/4-2). All remaining errors are our own.

<sup>†</sup>DIW Berlin, Potsdam University, ngohl@diw.de

<sup>‡</sup>DIW Berlin, FU Berlin, Netspar, phaen@diw.de

<sup>§</sup>DIW Berlin, CESifo, IZA, CEP/LSE, fweinhardt@diw.de

# 1 Introduction

Human capital theory starting with Ben-Porath (1967) and Becker (1962) predicts that the value of human capital investment increases with the payout period of the investment.<sup>1</sup> In this paper, we propose a novel empirical test of this key prediction of the human capital theory. We exploit an exogenous increase in the working life induced by a sizable pension reform which changes the payout period of the human capital investment. Our main finding based on theory and empirical analysis using a regression discontinuity design (RDD) is that an increase in the working life causally increases human capital investment.

This study presents new evidence that investments, in line with the human capital theory, respond to changes of the payout period. The policy-induced variation of the pension reform that we exploit affects the employed in the later part of their working career, i.e. after the age 50. This period is of central importance for the labor markets and employment policies since employment rates start to decline during that period. The reform is relevant for large parts of the working population so our results are informative for macroeconomic human capital theory as well as theory on pension design.

We proceed in two steps. In the first step, we theoretically discuss the implications of an increase in the working life on human capital accumulation. The model is based on the standard life cycle theory of human capital by Ben-Porath (1967) but we allow for human capital accumulation during the working life through on-the-job training (in the following training). Moreover, changes in the payout period for human capital investment are not determined by variation in life expectancy, but by a change in the working life. We derive three propositions from the model. Proposition I shows that investment in training increases with working life while according to Proposition II training is increasing with initial education. Finally, Proposition III determines the conditions when investments in training induced by an increase in the working life increase with initial schooling.

---

<sup>1</sup>This important prediction is the basis to explain the joint increases in life expectancy and educational investments that have been witnessed in most countries over the past century, see e.g. Soares (2005); Cervellati and Sunde (2013).

In the second step, we provide empirical evidence for the propositions of the life-cycle model. The empirical analysis is based on the data from the German microcensus; the microcensus is a representative yearly household survey which covers 1% of all German households (about 370,000 households per year). This data includes detailed information about specific job-related training which we use to measure post-schooling human capital investment. In the analysis, we exploit a pension reform to provide causal evidence for the effect of working life on human capital investment. Specifically, the pension reform abolished an important early retirement program for women born after 1951. Women born in 1951 and before could enter retirement at the age of 60 through this pathway. In contrast, for women born in 1952 and later this pathway has been closed; these women can enter retirement only at the age of 63. The reform is well suited to estimate the effect of an increase in working life on training and to test human capital theory. First, Geyer and Welteke (2019) document a sizable positive employment effect and a causal increase in working life induced by the pension reform. Second, the pension reform was already implemented in 1999. Thus, the affected women, i.e. women born in 1952, still had a long remaining working life to benefit from human capital investment. As a result, we can exploit the variation at the cut-off between these adjacent cohorts to isolate the causal effect of an increase in working life on human capital investment on the job.

In the empirical analysis we first use descriptive evidence from the microcensus to provide empirical support for Proposition I and II: In line with Proposition I, we show that training positively correlates with remaining working life. Moreover, training also positively correlates with initial educational levels (Proposition II). We then present causal support for Propositions I and III. Our empirical results show that the increase in the retirement age has a sizable effect on human capital accumulation: training increases depending on the specification by about 2.5-5 percentage points which corresponds to an increase of 20% - 30% for these age groups. We carefully examine robustness. Our key finding is robust to the omission or inclusion of individual-level control variables. Moreover, results are stable to changes in

the bandwidth for different specifications of the running variable and results do not change in Donut-RDD specifications that leave out the first or the first two month on each side of the cutoff. We also test for balancing in individual characteristics and show balance across the cutoff. Finally, we conduct placebo tests for which we artificially alter the cut-off date and document null effects of training.

Turning to heterogeneity, we find clear evidence that the training effect is increasing with initial education. The pension reform increases training for women with a college degree or more by about 13 percentage points, which corresponds to a relative increase of about 40%. The effect for women without college degree is not significant.

Our study is related to several strands of the literature. Most important, we contribute to empirical studies related to the human capital theory which estimate the effect of mortality on educational outcomes and economic growth, surveyed in e.g. Bloom et al. (2019). Studies using variation in mortality face at least two challenges. First, as discussed in Hazan (2009) and Cervellati and Sunde (2013), it is not the change in the length of life pre-se which matters for investment in human capital, but the survival rates during working life. Second, variation in life expectancy is rarely random or unexpected, complicating causal estimation. A large part of the empirical literature uses variation in mortality rates between countries or states, e.g. Acemoglu and Johnson (2007), Lorentzen et al. (2008) or Hansen and Strulik (2017), with mixed findings. Several papers specifically address the methodological challenges focusing on specific diseases, or on changes in health services in the context of developing countries. Oster et al. (2013) use variation in life expectancy driven by Huntington disease realisations across individuals who have ex-ante similar risks for realisations of this neurological disorder. They find effects in line with human capital theory on college attendance and completion, health outcomes, as well as on job training for individuals with different realisations or information (genetic testing) undertaken between the ages of 17 and 35.<sup>2</sup> In developing countries, Jayachandran and Lleras-Muney (2009) use a strong decline in maternal death

---

<sup>2</sup>In this study, job training is measured using a variable on whether individuals have ever undertaken job training for promotion or job advancement since starting their job.

rates in Sri Lanka and find positive effects on girls' educational investments measured in years of school education and literacy rates. In another important study, Baranov and Kohler (2018) exploit variation in mortality rates related to HIV medication in Malawi to study effects on savings and on children's educational investments. They find positive effects of an increase in life expectancy on both types of outcomes.

Our study is complementary to these studies as it presents evidence that is not related to variation in life expectancy and focuses on training in the later part of the working career instead of schooling as the central human capital variable. We use a regression discontinuity design and variation in working life induced by a pension reform, a common policy parameter across the developed world. This allows us to study effects of changes in working life on educational investment for a population that is relatively older and comprises an important part of the labor force.

The paper is also related to the literature which analyzes the effect of pension reforms on employment, income, training and mortality. In general, these studies document positive employment effects and an increase in the working life of pension reforms that reduce the generosity of the pension system. These studies either exploit exogenous variation in the pension rules for the identification<sup>3</sup> or they are based on structural retirement models<sup>4</sup>. Crucially, these studies typically assume an exogenous process of human capital investment which implies that individuals cannot adjust their human capital investment through additional training in response of a pension reform. Notable exception are the structural analyses by Fan et al. (2017) and Blundell et al. (2019). Fan et al. (2017) show that a reduction in the generosity of the pension system leads to an increase in human capital accumulation which is not consistent with the assumption of an exogenous human capital process and similarly, Blundell et al. (2019) document for women that human capital accumulation through training has positive effects which partly compensate the negative career effects of children.

---

<sup>3</sup>Examples include, Duggan et al. (2007), Mastrobuoni (2009), Staubli and Zweimueller (2013), Atalay and Barrett (2015), Manoli and Weber (2016), or, Geyer and Welteke (2019).

<sup>4</sup>See e.g. Gustman and Steinmeier (1986), Rust and Phelan (1997), French (2005), French and Jones (2011) or Haan and Prowse (2014)

Several papers also show that a reduction in the generosity of the pension system leads to an increase in training, e.g. Montizaan et al. (2010), Brunello and Comi (2013) or Bauer and Eichenberger (2017). However, these papers do not link their findings to a theoretical model and are based on smaller reforms or specific settings such as workers in large public sector firms only.<sup>5</sup> Finally, there is little empirical evidence that pension reforms have meaningful effects on life expectancy. Specifically, the literature shows mixed evidence about the effect of retirement on mortality, but in general the effects are relatively small, in particular for women. Kuhn et al. (2010) and Fitzpatrick and Moore (2018) document that retirement has a positive effect on mortality for men but no significant effect on female mortality. Hernaes et al. (2013) find that the retirement age has no effect on mortality.

The paper is organized as follows. In Section 2 we develop a model of human capital accumulation during the working life. Section 3 describes the German public pension system and the 1999 pension reform, introduces the data and provides descriptive support for the predictions of the human capital model. In Section 4, we turn to the causal analysis. We describe the method and document causal evidence from graphical and regression analyses. Finally, Section 5 concludes.

## 2 Theoretical model

In this section, we derive a theoretical human capital model and show that individuals, *ceteris paribus*, have an incentive to increase training when working life increases.<sup>6</sup> The central mechanism for this human capital effect is that the returns to training increase with the re-

---

<sup>5</sup>The paper is as well related to studies which focus on other dimensions of human capital investment and training. Several studies discuss the theory of human capital investment through training and provide empirical evidence about the effect on labor market outcomes in form of wages, job security or employment probability (see for example Pischke (2001), Zweimueller and Winter-Ebmer (2000), Barrett and O’Connell (2001), Leuven (2005), Frazis and Loewenstein (2005), Picchio and van Ours (2011), or Ruhose and Weilage (2019)).

<sup>6</sup>There exist several reasons why individuals have a general motivation to invest in training. Most important, empirical evidence shows that training has a positive effect on wages and on employment, see e.g. Frazis and Loewenstein (2005) and Blundell et al. (2019). Moreover, training can improve quality of work and can have positive effects on non-pecuniary outcomes Ruhose and Weilage (2019).

maintaining working life of an individual  $i$ , denoted by  $R_i$ .<sup>7</sup> Further, we analyze how the human capital effect induced by increasing working life differs by initial schooling,  $S_i$ . To derive testable predictions by initial schooling, it is necessary to establish first how training, *ceteris paribus*, varies with initial schooling endowment. With the assumption that initial schooling and further human capital accumulation are dynamic complements, see e.g. Cunha and Heckmann (2007) and Jacobs (2009), we can show that training increases with initial education. In the Appendix we show that we can replace this assumption of dynamic complementary by instead introducing utility costs of training as in Blundell et al. (2019). Finally, we determine the conditions when investments in training induced by an increase in the working life increase with initial schooling.

The theoretical model presented below illustrates the mechanism in a simplified and intuitive setting through a discrete time model consisting of three stylized periods. Note that  $Y_{ti}$  denotes an individual  $i$ 's income in period  $t$  and  $C_{ti}$  denotes the level of consumption in period  $t$ .

## 1. Period

Each individual derives utility through consumption,  $U(C_{1i})$ , with the standard assumption of  $U'(C_{1i}) > 0$ ,  $U''(C_{1i}) < 0$ .

Income in period one is given by:

$$Y_{1i} = w_1(S_i)(1 - I_i) \tag{1}$$

Each individual earns wage  $w_1(S_i)$ .<sup>8</sup> The wage in period one depends on the initial level

---

<sup>7</sup>An analogous human capital effect can be generated in a model of firm's investment decision when the working life of workers increases. When workers are not perfectly mobile (Acemoglu and Pischke, 1998, 1999), the intuition is straight forward in our model: The longer the payout period for the investment of the firm, i.e. the longer the worker stays in the firm, the higher the investment in human capital. In the data we can not observe if the training is initiated and financed by the worker or the firm. However, for the empirical test of the human capital effect this information is not required as we focus on the effect of working life on training investment and not on the incidence of the human capital investment.

<sup>8</sup>We assume that individuals are either full time employed or unemployed. Wages of the unemployed are zero. Thus an increase in the wage can result from entering employment or increase in earnings in full time employment.



of schooling which is determined exogenously prior to period one. The wage is increasing in education, specifically we assume  $w'_1(S_i) > 0$ ,  $w''_1(S_i) < 0$ . In this period the individual decides on his or her time investment in human capital,  $I_i$ , through participation in on-the-job training measures. Time investment in on-the-job training implies opportunity costs which in the model are characterized by a fraction of foregone labour income,  $w_1(S_i)I_i$ . Hence, individuals with high education face higher opportunity costs.

## 2. Period

Income in period two is given by:

$$Y_{2i} = w_2(S_i, I_i)R_i \quad (2)$$

$R_i$  is the duration of the remaining working life and  $w_2(S_i, I_i)$  is the wage earned in period two. Individuals in period two collect the returns from their human capital investment made in period one through their wage, as the wage in period two is a function of training as well as initial schooling, i.e.  $w_2(I_i, S_i)$ . Note that by assumption  $w_{2I} > 0$  and  $w_{2II} < 0$ . Further, in order to capture the dynamic complementarity between initial schooling,  $S_i$  and human capital investment,  $I_i$ , we assume  $w_{2IS} > 0$ .

## 3. Period

Period three is the period of retirement. The duration of period three is  $T_i - R_i$ , where  $T_i$  is the individual life expectancy. We assume that retirement is a discrete decision to exit the labor market completely. Income in the retirement period is covered by the state pension which is a fraction  $\alpha$ , with  $\alpha < 1$ , of labour income in period two.

$$Y_3 = \alpha w_2(S_i, I_i)(T_i - R_i) \quad (3)$$

Utility over all three periods is given by:

$$U_G = U(C_{1i}) + \beta U(C_{2i})R_i + \beta^2 U(C_{3i})(T_i - R_i)$$

where  $\beta$  is the discount factors, with  $\beta < 1$ . The intertemporal budget constraint is:

$$y_{1i} + \beta y_{2i} + \beta^2 y_{3i} \geq C_{1i} + \beta C_{2i}R_i + \beta^2 C_{3i}(T_i - R_i)$$

The maximization problem can hence be depicted by the Lagrangian:

$$L = U_G + \lambda [y_{1i} + \beta y_{2i} + \beta^2 y_{3i} - (C_{1i} + \beta C_{2i}R_i + \beta^2 C_{3i}(T_i - R_i))] \quad (4)$$

The details of the solutions and the set of First Order Conditions are presented in the Appendix.

Using comparative statics we derive three main propositions from the human capital model. Firstly, in Equation 5 we show how the effect of an exogenous increase in the remaining working life duration ( $R_i$ ), affects the investment in training ( $I_i$ ).

$$\frac{\partial I_i}{\partial R_i} = \frac{[\alpha\beta - 1] w_{2I}(S_i, I_i)}{[R_i + (T_i - R_i)\alpha\beta] w_{2II}(S_i, I_i)} \quad (5)$$

By assumption  $w_{2I} > 0$  and  $w_{2II} < 0$ . Further,  $\alpha\beta - 1$  is negative<sup>9</sup>. Therefore the expression above is positive and we can derive the following proposition.

**Proposition I (Working life effect)**

*The effect of an increase in the working life on training is positive.*

Further, we discuss how initial schooling affects the amount of time allocated towards on-the-job training. The relation between initial schooling and training is captured by Equation 6.<sup>10</sup>

---

<sup>9</sup>This follows from  $\alpha < 1$  and  $\beta < 1$ .

<sup>10</sup>For a detailed derivation of Equation 6 see Appendix

$$\frac{\partial I_i}{\partial S_i} \geq 0 \iff w_{2IS}(S_i, I_i) \frac{S_i}{w_{2I}(S_i, I_i)} \geq w_{1S} \frac{S_i}{w_1(S_i)} \quad (6)$$

The above shows that as long as the relative return to human capital investment collected in period two,  $w_{2IS}(S_i, I_i)$ , proportionally out-weighs the opportunity costs, i.e. the return of schooling received in period one,  $w_{1S}$ , human capital investment will increase in schooling, i.e.  $\frac{\partial I_i}{\partial S_i} > 0$ .

**Proposition II (Initial education effect)**

*The level of time investment in training rises with initial schooling as long as the additional returns to human capital investment related to the initial level of schooling are larger than the opportunity cost of the investment*

Finally, in Equation 7 we discuss how the effect of the increase in the working life on training varies by the initial level of schooling. This effect can be decomposed in two effects.

$$\frac{\partial}{\partial S_i} \left[ \frac{\partial I_i}{\partial R_i} \right] = \frac{\partial I_i^2}{\partial R_i \partial S_i} + \frac{\partial I_i^2}{\partial R_i \partial I_i} \frac{\partial I_i}{\partial S_i} \quad (7)$$

The first effect captures the direct effect of how the working life effect,  $\frac{\partial I_i}{\partial R_i}$  (Proposition I), changes with initial schooling and the second effect captures the indirect effect of how the working life effect is influenced by the existing level of training given the initial schooling effect  $\frac{\partial I_i}{\partial S_i}$  (Proposition II). Under the assumption of complementarity between schooling and training (or alternatively about the utility cost of training, see Appendix) the direct effect is positive. Therefore, individuals with a higher initial level of schooling have a stronger incentive to invest in training as a response to an increase in the working life. The indirect effect, however, is negative as the returns to training are positive but decreasing. Given an increase in the working life, individuals with higher levels of training, therefore, have a lower incentive to further increase their training investment. The overall effect ultimately depends on the two opposing sub-effects, as summarized in Proposition III. In the Appendix we present a detailed formal derivation of the two opposing effects and their respective signs.

**Proposition III (Working life effect by education)**

*The reform effect rises with the initial level of schooling, as long as the positive direct effect outweighs the negative indirect effect.*

The remainder of the paper will empirically assess the predictions derived from the theoretical model depicted above. First we will present descriptive evidence to provide empirical support for Propositions I and II. Then, we exploit the variation in the working life induced by the pension reform to causally test Propositions I and III.

### 3 Institutional Setting and Data

#### 3.1 Pension reform

Before we turn to the empirical analysis we summarize the relevant aspects about the German pension system and the 1999 pension reform which induces exogenous variation in the working life.

The statutory public pension system is the central part of the pension system in Germany. It covers more than 80% of the workforce with the exceptions of groups that are not subject to compulsory pension insurance, most important civil servants, and self-employed. It includes old-age pensions, disability pensions, and survivors benefits. The system is financed by a pay-as-you-go (PAYG) scheme and has a strong contributory link: pension benefits depend on the entire working history. The pension system provides several pathways into early retirement, i.e. claiming retirement benefits before reaching the normal retirement age. In this analysis we focus on the *pension for women* which allows drawing benefits starting from age 60.<sup>11</sup>

The 1999 reform abolished the *pension for women* for cohorts born after 1951. Effectively, the reform raised the early retirement age (ERA) for most women from age 60 to age 63 and

---

<sup>11</sup>In addition early retirement is possible via : (1) the *invalidity pension*; (2) the *pension after unemployment or after old-age part-time work*; (3) the *pension for the long-term insured*, for more details see Geyer et al. (2018). For a more general description on the German pension system, see Boersch-Supan and Wilke (2004).

therefore increased the working life.<sup>12</sup> Women born before 1952 could claim the *pension for women* if they fulfilled certain qualifying conditions. The eligibility criteria were: (i) at least 15 years of pension insurance contributions; and (ii) at least 10 years of pension insurance contributions. According to Geyer and Welteke (2019), about 60% of all women born in 1951 were eligible for the old-age pension for women. In our empirical analysis we focus only on employed women, about 75% of these women fulfill the criteria and are therefore eligibility for this pathway (see Table 6 below). The pension reform was implemented when affected women born in 1952 were aged 47. Thus, these women had still a long horizon to benefit from human capital investments.

Geyer and Welteke (2019) and Geyer et al. (2018) have evaluated the labor market effects of the pension reform based on administrative data of the public pension insurance accounts and the Microcensus, respectively. Several findings of these studies are relevant for the following empirical analysis. Most important the increase in the ERA has a sizable positive effect on the working life on individuals which is the necessary condition for an increase in human capital investment. In more detail, employment rates for eligible women aged between 60 and 62 increase by about 15 percentage points, the combined effect on inactivity and unemployment has with about 12 percentage points a similar size. Second, Geyer and Welteke (2019) document that the pension reform has no significant effect on employment, unemployment, disability or inactivity before the age of 60. This implies, estimation results on the effect on human capital accumulation for employed women before the age of 60 are not affected by selection effects into employment induced by the pension reform. Third, there exists effect heterogeneity by initial education (Geyer et al., 2018). Specifically, employment effects for women aged 60-62 are larger for high educated women (9.5%) but they are still sizable and significant for women without high education (8.2%). We will return to this finding when we discuss heterogeneity in section 4.7.

---

<sup>12</sup>The *pension after unemployment or after old-age part-time work* was abolished at the same time as the *pension for women*. However, this does not affect our analysis, as the ERA for this pension type was already 63.

### 3.2 The German Microcensus

For the analysis we use the German Microcensus. The Microcensus is an annual, household based survey with representative information about the population and the labour market in Germany. Participation in the survey is mandatory. It has a sampling fraction of one percent of the German population (about 370 000 households) and constitutes the largest annual household survey in Europe (RDC of the Federal Statistical Office and Statistical Offices of the Laender, 2015).

In the main analysis we concentrate on employed<sup>13</sup> women younger than 60 years born in 1951 and 1952 which we observe during the period 2005 to 2012. For these years the data include information about the month of birth and consistent information about the participation of on-the-job training.<sup>14</sup> We observe around 1,250 individuals for each birth month in our sample. Thus, overall, the sample includes information of about 30,000 women born in the two cohorts of interest. The Microcensus includes important socio-demographic variables, such as age, education,<sup>15</sup> marital status or firm size which we use as control variables.

### 3.3 On the job training: descriptive evidence

The Microcensus provides in addition information if an employed person has participated in on-the-job training during the last twelve months. The training information includes specifically courses that are related to career development, e.g. to improve management, computer or rhetoric skills.<sup>16</sup>

In the following we describe how training participation varies by age and education which

---

<sup>13</sup>Women working in "mini-jobs" are not counted as employed.

<sup>14</sup>Before the year 2005 the Microcensus only provides information about the birth year and the definition of training changes at several points in time. Therefore, the extension of the sample before 2005 would require additional assumptions.

<sup>15</sup>Education is measured with ISCED 2011 levels: based on this information with define women without college degree or with college degree or more.

<sup>16</sup>The exact wording of the question reads: *Did you in the last 12 months take part in any form of vocational training? Examples of vocational training are: occupational re-training, courses for career development and general training courses for example in the fields computing, management, public speaking.*

allows us to provide descriptive evidence for Propositions I and II.<sup>17</sup> Figure 1 shows the age pattern of training for all employed women born between 1940 and 1997. As expected we find a hump shaped age profile. Training rates are low at very young ages however they rapidly increase in the first years of the working career. Towards the end of the working life, i.e., after the age of 50, training strongly decline. Specifically, Figure 1 suggests that training rates decline between age 53 and 58 by about 10 percentage points. Hence, the period after the age of 50, which we study in the causal analysis, is particularly sensitive to changes in the incentives for training. This implies even small changes in the incentives to invest into training can lead to large responses. In summary, the pattern of Figure 1 provides first support for Proposition I. Training is reduced when the working horizon, i.e. the payout period, decreases.

[Figure 1 about here]

In Figure 2 we turn to the training pattern by initial education. In line with Proposition II we find that training increases with the level of education. Specifically, employed women born in 1951 with no college degree ( $\text{ISCED} < 5$ ) have training rates of about 15.4 percent. In contrast women with a college degree or higher tertiary education ( $\text{ISCED} > 5$ ) have training rates of 31.7 percent.

[Figure 2 about here]

## 4 Causal Analysis: RDD

### 4.1 Empirical method

In the empirical analysis we exploit the 1999 pension reform to estimate the effect of an increase in working life on human capital investment and to provide a causal test of Propositions

---

<sup>17</sup>For more information on training participation and its measurement in Germany see for example Eiser-mann et al. (2014).

I and III using a RDD. The reform leads to an arbitrary and distinct cutoff for women born before and after December 31, 1951, which determines the assignment into the treatment and the control group.

More formally, in the empirical analysis the woman’s month of birth is the running variable  $M$ , which determines treatment  $D$  as one if she was born after December 31, 1951 and zero otherwise:

$$D_i = \begin{cases} 1, & \text{if } M_i \geq c \\ 0, & \text{if } M_i < c \end{cases} \quad (8)$$

For the identification of a causal effect it is important that no manipulation of the month of birth for women born in 1951 and 1952 and no selection into or out-of treatment is possible. As a result, the treatment and control groups should be otherwise comparable around the cut-off. We provide supporting evidence based on balancing tests of important pre-policy covariates of the 1951 and 1952 birth cohorts, and by moving the cutoff to hypothetical placebo dates. Moreover, as discussed e.g. in Geyer and Welteke (2019) no other relevant policy reform differently affected women born in 1951 and 1952.

In the main specification we implement the RDD in the following regression model:

$$y_i = \alpha + \beta D_i + \gamma_0 f(M_i - c) + \gamma_1 D_i f(M_i - c) + X_i \delta + \varepsilon_{it} \quad (9)$$

$D_i$  is a dummy specifying treatment, that is equal to 1 if a woman is born 1.1.1952 or later and 0 otherwise. A woman’s month of birth is described by  $M_i$  and  $c$  is the cut-off date for the increase in early retirement age, ERA (January 1, 1952). Therefore,  $f$  is a function of the difference between a woman’s birth date and the beginning of the ERA increase  $M_i - c$ , i.e. the running variable. This function is interacted with the treatment variable  $D_i$  to allow for different slopes before and after the cutoff. In addition we account for further explanatory variables ( $X$ ), including age, education, marital status, firm size and regional information. We implement this specifications using either global polynomials for  $f$  up to the third degree,



as well as local polynomials, and for various bandwidth choices.<sup>18</sup>

The outcome variable  $Y$  in our analysis is on-the-job training which is dichotomous i.e. taking on the value 1 if a women has participated in training in the last twelve months and 0 if she has not. In the main specification we implement the RDD in a linear probability model.<sup>19</sup>

## 4.2 Graphical analysis

Before turning to the estimation result we provide graphical evidence how an increase in the working life affects training. In Figure 3, we present the pattern of training for the relevant cohorts, i.e women born 1951 and 1952, for the ages which we can observe in the data. The Figure clearly shows that women with a longer working horizon, i.e. women of cohort 1952, have higher training rates than women born in 1951. As expected, the difference gets smaller close to age 60, i.e. when both groups have only a short working horizon.

[Figure 3 about here]

In the next Figure we present how training participation varies around the cut-off date. Specifically, Figure 4 shows participation rates in training by month of birth 12 months before and after the cut-off birth date, 1.1.1952. The share of employed women participating in training is clearly higher after the cut-off. The average rate of participation in the 12 month before the cut-off date is approximately 15.4 %. After the cut-off date, the graphs shows a jump in the average rate of training participation for employed women under 60 to more than 16.5 percent. In the next section, we examine the robustness and significance of the the graphical evidence using various choices in the RDD framework.

[Figure 4 about here]

---

<sup>18</sup>Local polynomials are estimated using the Stata package "rdrobust" (Calonico et al., 2018).

<sup>19</sup>Estimation results based of a Probit model (not reported) show very similar results.

### 4.3 Regression results - overall effects

To quantify the effect of an increase in the working life on the investment into human capital we use the RDD described in Section 4.1. In the analysis, we focus on all employed women in their later working life, i.e. when they are aged between 53 and 60.

In Table 1 we present the estimation results for different specifications with observations 12 months before and after the cut-off date. We consider regressions with global polynomials, with linear, quadratic and cubic specifications and local linear and local quadratic specifications. Moreover, the table includes these regressions without and with additional control variables. Standard errors are reported in brackets and clustered at the birth month level. Our inference is robust to a specification without clustered standard errors as suggested by Kolesar and Rothe (2018).

[Table 1 about here]

The results of these different specifications all point in the same direction despite some expected differences in the magnitude of the point estimates: the increase in the early retirement age has a positive and significant effect on the investment in training. Most point estimates show that the participation in training increases between 2.5-5 percentage points, which translates into a relative increase of about 20-30% given the pre-reform share in training of 15.4%.

Only the global linear specification in the top panel (without covariates) in column 1 is not statistically significant, albeit positive. However, this is the least flexible version of conditioning on the running variable and therefore -a priori- not the preferred specification. In contrast, the local linear regressions in Columns 4 to 5 of Table 1 consistently show positive and significant estimates in similar magnitude across all specifications.

### 4.4 Balancing tests

[Table 2 about here]

The assumption that underlies the RD design is that other factors vary smoothly across the cutoff. We provide support for this assumption by using individual control variables as outcomes using the same specification as in Table 1. The resulting estimates are presented in Table 2. All control variables, except the share of women with high income, are insignificant at conventional levels of statistical significance. Yet, the income variable is significant only at the 10% level of statistical significance. The balancing in these variables is also reflected in the fact that adding the controls to our main specification hardly affects the estimates for the training outcome, as already seen in Table 1.

## 4.5 Placebo analysis

[Table 3 about here]

In addition, we conduct two placebo analyses presented in Table 3. In the first placebo analysis we artificially shift the cut-off date by one year to 1.1.1950 and in the second placebo analysis to 1.1.1952. Importantly, the pension rules are identical before and after the chosen placebo cut-offs. The shift of one entire year is of particular relevance as this could capture potential seasonal effects related to the December to January timing of the reform introduction.

The result from this additional analysis support our identification strategy: the treatment effect is very close to zero and not significant in both placebo specifications, with and without additional explanatory variables. Moreover, these effects are precisely estimated clearly differ from our main findings in Table 1.

## 4.6 Further robustness checks: bandwidth and donut regressions

Bandwidth choices can affect RD estimates so we carefully examine if and how our main results are sensitive to these. First, in Appendix Table A.1 we replicate our main results from Table 1 but using a bandwidth of six month rather than a full year on both sides of

the cutoff, for all specifications of the running variable. Second, in Appendix Table A.2 we show that our results also hold for additional bandwidth choices for local polynomials, where we present result for bandwidth choices of 6, 9 and 12 month. The bandwidth of 6 month is the chosen bandwidth of the endogenous bandwidth selection routine "rdbwselect", using the mean squared error criterion and a triangular kernel (Calonico et al., 2014). In all cases our estimates remain in the same ballpark. All estimates except the second column in Appendix Table A.1, which shows results of the global analysis with a quadratic polynomial, remain statistically significant at the ten percent level or higher.

As additional robustness check, we also ensure that observations close to the cutoff do not cause our results. Specifically, we examine if observations close the to cutoff drive our effects by estimating effects from donut-RD regressions. We estimate different specification of Equation 9 without one or two birth month closest to the cutoff on both sides. Appendix Table A.3 shows the resulting estimates for the various functional form choices and without and with individual control variables. Some of the specifications without covariates loose statistical significance in the one-month donut, presented in the upper panel. However, overall this additional analysis confirms the main findings, with estimates of similar magnitudes throughout. Our results are not driven by observations close to the cutoff.

## 4.7 Effects by initial education

We now extend the empirical analysis and focus on effect heterogeneity along prior educational levels. We have already shown descriptively that training participation positively correlates with prior educational levels in Section 3.3 (Proposition II). We now move on to test if the reform effect also varies by prior educational level (Proposition III). Since we only focus on women close to the cut-off, to alleviate issues of sample size we aggregate the ISCED educational levels into two education groups of women with "college" and "non-college".

[Table 4 about here]

We find very strong differences by education for the different specifications presented in Table 4. Women with college education or more increase training by nearly 12 percentage points which corresponds to a relative increase of about 35%. The effect for women without college education is estimated to be close to zero and not significant at conventional levels. We therefore conclude that the direct training effects outweighs the negative indirect effect (Proposition III), which explains why the human capital induced by an increase in working life increase with education.

## 4.8 Quantification of effects

As discussed in Section 3.1, the 1999 pension reform only affected the working life of women which fulfill the eligibility criteria for the so-called pension *for women*. The Microcensus, which we base our analysis on, is a cross sectional data set without information about the employment history. Therefore, we can not directly determine the eligibility within this data. As a result, our estimates should be interpreted as "intention-to-treat" (ITT) effects, giving a lower bound of the true effect.

To gauge information about actual eligibility, we use information from the longitudinal data of the Socio-economics panel (SOEP), the largest household panel data set available in Germany. In the representative SOEP data, about 76% of women employed before entering retirement were indeed eligible for this pathway into retirement. Further, the SOEP data show that about 72% of employment women without a college degree and 85% of women with a college degree fulfill the eligibility criteria..

With this information and the estimated effects (ITT) presented in Tables 1 and 4, we can derive the average treatment effect on the treated (ATT) for the different educational groups. These are presented in Table 5. Overall, the pattern of the ATT effects is similar to the ITT effects, but the effects are slightly larger. Importantly, the ATT effects provide as well empirical support for Propositions I and III derived from the human capital model. The point estimates suggest that overall women training increases by 4.6 percentage points

(column 1), for women with college the increase is over 13 percentage points (column 2) and the effect for women without college is close to zero (column 3). The point estimates imply a relative increase in training of 30% of all women and 40% of women with college degree.

[Table 5 about here]

Overall, the magnitude of the estimated ITT and ATT effects are consistent with the strongly declining pattern of training after the age of 50, presented in Figure 1. As discussed above, training participation is particularly sensitive to changes in ages close to retirement which determine the remaining working life. Therefore, it is not surprising that an increase in the working life induced by the increase in retirement age from 60 to 63 has sizable effects on training.

## 5 Conclusion

In this paper we provide causal evidence for the theory of human capital accumulation. In the empirical analysis we use an exogenous change in the working life which increases the payout period for the human capital investment. Specifically we exploit a sizable pension reform which increased the early retirement age for women between two adjacent cohorts from 60 to 63 years. The analysis is based on the German Microcensus and exploits in a RDD the cohort specific variation of the pension reform.

The empirical analysis offers support for three key propositions of the human capital model which are derived in the first part of the paper. First we provide descriptive evidence that training positively correlates with remaining working life and with initial educational levels. We then present causal evidence that an increase in working life induced by the pension reform has a positive effect on human capital investment and that this human capital effect increases with initial schooling. In more detail, our empirical results show that the increase in the retirement age has a sizable effect on human capital accumulation of employed women aged 63-60: depending on the specification training increases by about 2.5-5 percentage

points which corresponds to an increase of 20-30% for these age groups. This finding is robust to changes in the bandwidth and for different specifications of the running variable in the RDD and is supported by placebo test. Moreover, we show that the pension reform increases training for women with a college degree or more by 11 percentage points, which corresponds to a relative increase of about 35%. The effect for women without college degree is not significant.

Besides testing a key prediction of human capital theory for a large and relevant part of the working population, our results have important implications for the policy debate about pension reforms. This debate usually abstracts from the dynamic human capital investment that we document. Future work should examine the role of individual workers and firms in initiating the positive training effects that we document, adding to the still relatively underdeveloped literature on educational investments beyond initial schooling.

## References

- ACEMOGLU, D. AND S. JOHNSON (2007): “Disease and development: The effect of life expectancy on economic growth,” *Journal of Political Economy*, 115, 925–985.
- ACEMOGLU, D. AND J. PISCHKE (1999): “The Structure of Wages and Investment in General Training,” *Journal of Political Economy*, 107, 539–572.
- ACEMOGLU, D. AND J.-S. PISCHKE (1998): “Why Do Firms Train? Theory and Evidence,” *The Quarterly Journal of Economics*, 113, 79–119.
- ATALAY, K. AND G. F. BARRETT (2015): “The impact of age pension eligibility age on retirement and program dependence: Evidence from an Australian experiment,” *Review of Economics and Statistics*, 97, 71–87.
- BARANOV, V. AND H.-P. KOHLER (2018): “The Impact of AIDS Treatment on Savings and Human Capital Investment in Malawi,” *American Economic Journal: Applied Economics*, 10, 266–306.
- BARRETT, A. AND P. J. O’CONNELL (2001): “Does Training Generally Work? The Returns to in-Company Training,” *ILR Review*, 54, 647–662.
- BAUER, A. B. AND R. EICHENBERGER (2017): “Endogenous aging: How statutory retirement age drives human and social capital,” CREMA Working Paper Series 2017-02, Center for Research in Economics, Management and the Arts.
- BECKER, G. (1962): “Investment in Human Capital: A Theoretical Analysis,” *Journal of Political Economy*, 70.
- BEN-PORATH, Y. (1967): “The Production of Human Capital and the Life Cycle of Earnings,” *Journal of Political Economy*, 75, 352–365.
- BLOOM, D., M. KUHN, AND K. PRETTNER (2019): “Health and growth,” *Oxford Research Encyclopedia on Economics and Finance*.
- BLUNDELL, R., M. C. DIAS, D. A. GOLL, AND C. MEGHIR (2019): “Wages, Experience and Training of Women over the Lifecycle,” *NBER Working Paper*, 25776.
- BOERSCH-SUPAN, A. AND C. WILKE (2004): “The German Public Pension System: How it Was, How it Will Be,” *NBER Working Paper* 10525.
- BRUNELLO, G. AND S. COMI (2013): “The Side Effect of Pension Reforms on Training: Evidence from Italy,” IZA Discussion Papers 7755, Institute for the Study of Labor (IZA).



- CALONICO, S., M. D. CATTANEO, M. H. FARRELL, AND R. TITIUNIK (2018): “RDROBUST: Stata module to provide robust data-driven inference in the regression-discontinuity design,” Statistical Software Components, Boston College Department of Economics.
- CALONICO, S., M. D. CATTANEO, AND R. TITIUNIK (2014): “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs,” *Econometrica*, 82, 2295–2326.
- CERVELLATI, M. AND U. SUNDE (2013): “Life Expectancy, Schooling, and Lifetime Labor Supply: Theory and Evidence Revisited,” *Econometrica*, 81, 2055–2086.
- CUNHA, F. AND J. HECKMANN (2007): “The Technology of Skill Formation,” *American Economic Review*, 2, 31–47.
- DUGGAN, M., P. SINGLETON, AND J. SONG (2007): “Aching to retire? The rise in the full retirement age and its impact on the social security disability rolls,” *Journal of Public Economics*, 91, 1327–1350.
- EISERMANN, M., F. JANIK, AND T. KRUPPE (2014): “Weiterbildungsbeteiligungursachen unterschiedlicher Teilnahmequoten in verschiedenen Datenquellen.” *Zeitschrift für Erziehungswissenschaft*, 17, 473–495.
- FAN, X., A. SESHADRI, AND C. TABER (2017): “Understanding Earnings, Labor Supply, and Retirement Decisions,” *Michigan Retirement Research Center Working Paper 367*.
- FITZPATRICK, M. D. AND T. J. MOORE (2018): “The Mortality Effects of Retirement: Evidence from Social Security Eligibility at Age 62.” *Journal of Public Economics*, 157, 121 – 137.
- FRAZIS, H. AND M. A. LOEWENSTEIN (2005): “Reexamining the Returns to Training: Functional Form, Magnitude, and Interpretation,” *Journal of Human Resource*, 40, 453–476.
- FRENCH, E. (2005): “The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour,” *The Review of Economic Studies*, 72, pp. 395–427.
- FRENCH, E. AND J. B. JONES (2011): “The Effects of Health Insurance and Self-Insurance on Retirement Behavior,” *Econometrica*, 79, 693–732.
- GEYER, J., P. HAAN, A. HAMMERSCHMID, AND M. PETERS (2018): “Labor Market and Distributional Effects of an Increase in the Retirement Age,” *IZA Discussion Paper 11618*.

- GEYER, J. AND C. WELTEKE (2019): “Closing Routes to Retirement: How Do People Respond?” *Journal of Human Resources*.
- GUSTMAN, A. L. AND T. L. STEINMEIER (1986): “A Structural Retirement Model,” *Econometrica*, 54, 555–584.
- HAAN, P. AND V. PROWSE (2014): “Longevity, life-cycle behavior and pension reform ,” *Journal of Econometrics*, 178, Part 3, 582 – 601.
- HANSEN, C. AND H. STRULIK (2017): “Life expectancy and education: evidence from the cardiovascular revolution,” *Journal of Economic Growth*, 22, 421–450.
- HAZAN, M. (2009): “Longevity and Lifetime Labor Supply: Evidence and Implications,” *Econometrica*, 77, 1829–1863.
- HERNAES, E., S. MARKUSSEN, J. PIGGOTT, AND O. L. VESTAD (2013): “Does Retirement Age Impact Mortality?” *Journal of Health Economics*, 32, 586–598.
- JACOBS, B. (2009): “Human Capital, Retirement and Pension Saving,” *ESF Forward Looks - Ageing, Health and Pensions in Europe*, 1–41.
- JAYACHANDRAN, S. AND A. LLERAS-MUNEY (2009): “Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines,” *The Quarterly Journal of Economics*, 124, 349–397.
- KOLESAR, M. AND C. ROTHE (2018): “Inference in Regression Discontinuity Designs with a Discrete Running Variable,” *American Economic Review*, 108, 2277–2304.
- KUHN, A., J.-P. WUELLRICH, AND J. ZWEIMUELLER (2010): “Fatal Attraction? Access to Early Retirement and Mortality.” *IZA Discussion Papers 5160*,.
- LEUVEN, E. (2005): “The Economics of Private Sector Training: A Survey of the Literature,” *Journal of Economic Surveys*, 19, 91–111.
- LORENTZEN, P., J. MCMILLAN, AND R. WACZIARG (2008): “Death and development,” *Journal of Economic Growth*, 13, 81–124.
- MANOLI, D. S. AND A. WEBER (2016): “The Effects of the Early Retirement Age on Retirement Decisions,” Working Paper 22561, National Bureau of Economic Research.
- MASTROBUONI, G. (2009): “Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities,” *Journal of Public Economics*, 93, 1224–1233.

- MONTIZAAN, R., F. COERVERS, AND A. DE GRIP (2010): “The effects of pension rights and retirement age on training participation: Evidence from a natural experiment,” *Labour Economics*, 17, 240–247.
- OSTER, E., I. SHOULSON, AND E. R. DORSEY (2013): “Limited Life Expectancy, Human Capital and Health Investments,” *American Economic Review*, 103, 1977–2002.
- PICCHIO, M. AND J. C. VAN OURS (2011): “Retaining Through Training; Even for Older Workers,” Tech. rep., SSRN Electronic Journal.
- PISCHKE, J.-S. (2001): “Continuous Training in Germany,” Tech. Rep. 3, Journal of Population Economics.
- RDC OF THE FEDERAL STATISTICAL OFFICE AND STATISTICAL OFFICES OF THE LAENDER (2015): “Microcensus, survey years 1996-2015, own calculations,” .
- RUHOSE, JENS, A. S. L. T. AND I. WEILAGE (2019): “The Benefits of Adult Learning: Work-Related Training, Social Capital, and Earnings,” *Economics of Education Review*, 72, 166–186.
- RUST, J. AND C. PHELAN (1997): “How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets,” *Econometrica*, 65, 781–832.
- SOARES, R. R. (2005): “Mortality Reductions, Educational Attainment, and Fertility Choice,” *The American Economic Review*, 95, 580–601.
- STAUBLI, S. AND J. ZWEIMUELLER (2013): “Does raising the early retirement age increase employment of older workers?” *Journal of Public Economics*, 108, 17–32.
- ZWEIMUELLER, J. AND R. WINTER-EBMER (2000): “Firm-specific Training: Consequences for Job Mobility,” IZA Discussion Papers 138, Institute for the Study of Labor (IZA).

# Figures and Tables

Figure 1: Average on-the-job training participation by age, Cohorts: 1940-1997

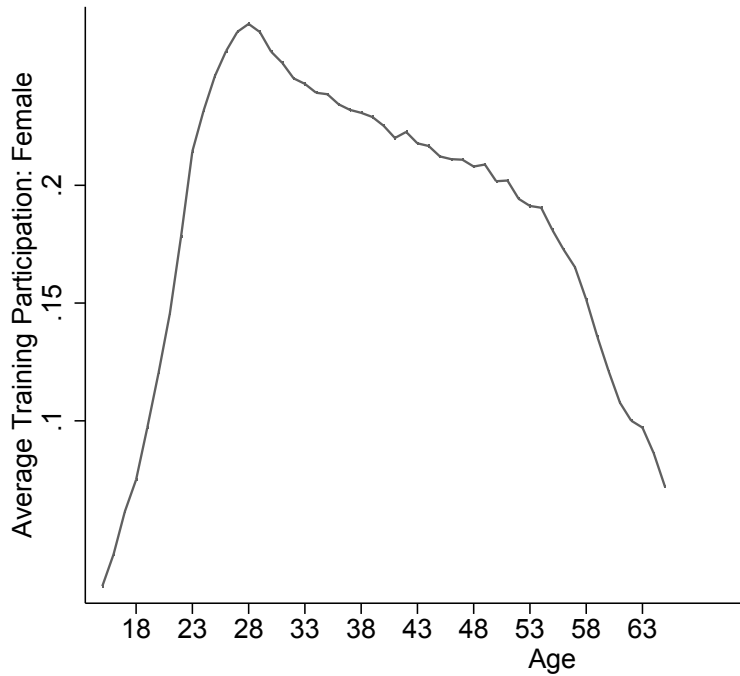


Figure 2: Average on-the-job training participation by ISCED groups

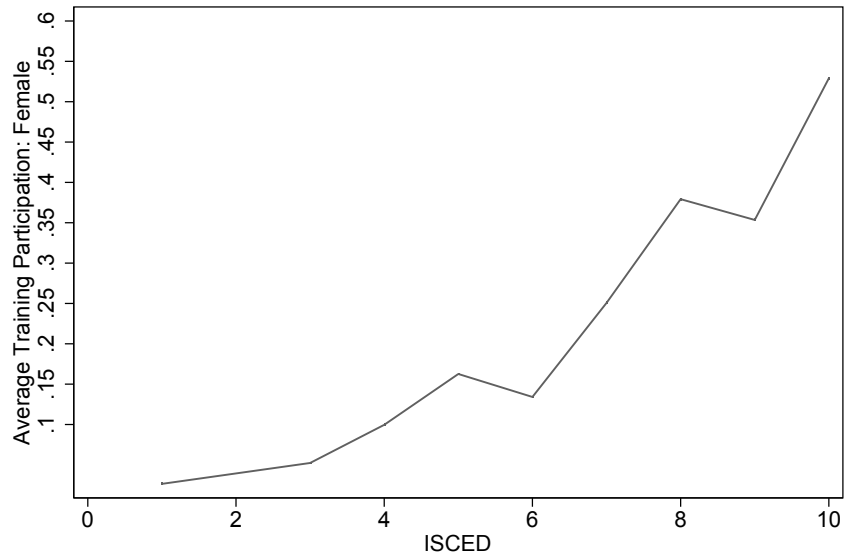


Figure 3: Average on-the-job training participation for sample group

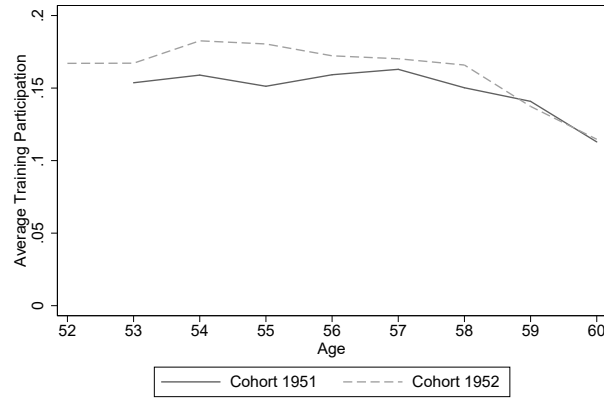
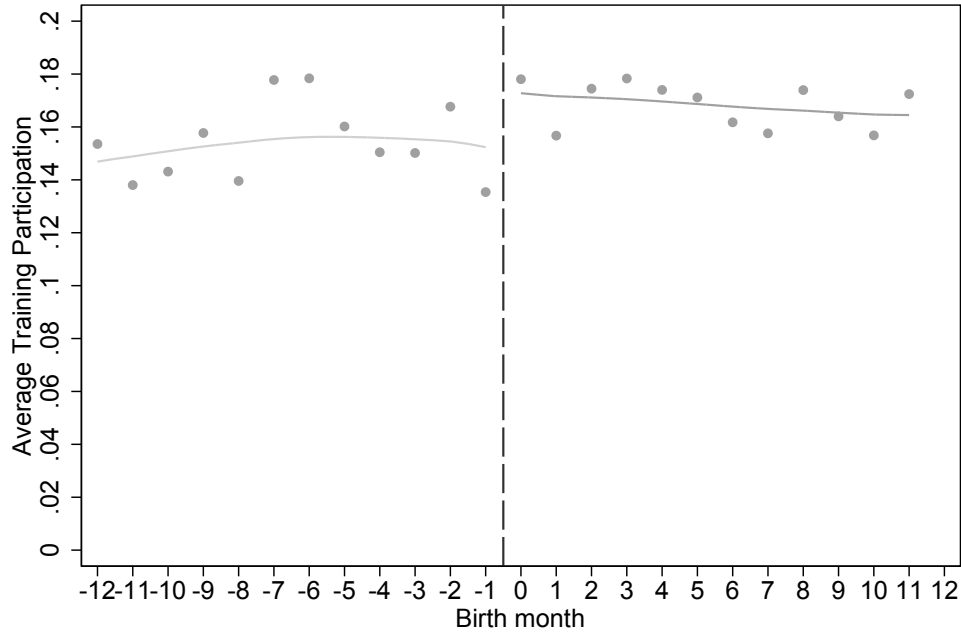


Figure 4: On the job training around the cut-off date



*Notes:* The fitted lines are local linear regressions using a first degree polynomial, a triangular kernel. In total, information for 13,658 individuals below the threshold and 14,873 individuals above the threshold are used.

Table 1: Regression Discontinuity: Main Results

	Global Polynomial			Local Polynomial	
	(1)	(2)	(3)	(4)	(5)
<b>Without Covariates</b>					
Treatment Variable	0.0157 (0.0147)	0.0352*** (0.0142)	0.0502*** (0.0105)	0.0235** (0.012)	0.0418*** (0.0131)
<b>With Covariates</b>					
Treatment Variable	0.0150* (0.0082)	0.0358*** (0.0104)	0.0439*** (0.0110)	0.0252*** (0.0087)	0.0413*** (0.0097)
Running Variable	Linear	Quadratic	Cubic	Linear	Quadratic

*Notes:* Standard Errors in parentheses; clustered at birth month level. Sample includes twelve month before and after reform cutoff. Significance levels: \* p 0.10, \*\* p 0.05, \*\*\* p 0.01; Pre-Policy Mean: 15.41 percent. Source: Microcensus 2005-2012, own calculations.

Table 2: Balancing

	Global Polynomial			Local Polynomial	
	(1)	(2)	(3)	(4)	(5)
<b>College Education</b>					
Treatment Variable	-0.0010 (0.0175)	0.0087 (0.0300)	0.0324 (0.0322)	0.0029 (0.0230)	0.0191 (0.0307)
<b>High Household Income</b>					
Treatment Variable	-0.0168 (0.0108)	-0.0295* (0.0142)	0.0040 (0.0168)	-0.0220* (0.0122)	-0.0149 (0.0186)
<b>West</b>					
Treatment Variable	0.0156 (0.0192)	0.0056 (0.0297)	-0.0097 (0.0415)	0.01162 (0.0233)	-0.0010 (0.0326)
<b>Big Company</b>					
Treatment Variable	-0.0061 (0.0096)	-0.0129 (0.0121)	0.0221 (0.0213)	-0.00900 (0.0080)	0.0023 (0.0104)
<b>Single</b>					
Treatment Variable	-0.0093 (0.0058)	-0.0083 (0.0065)	-0.0006 (0.0089)	-0.0089* (0.0051)	-0.0050 (0.0063)
Running Variable	Linear	Quadratic	Cubic	Linear	Quadratic

*Notes:* Standard Errors in parentheses. Sample includes twelve month before and after reform cutoff. Significance levels: \* p 0.10, \*\* p 0.05, \*\*\* p 0.01; Source: Microcensus 2005-2012, own calculations.

Table 3: Placebo Regression Discontinuity Designs

	Global Polynomial			Local Polynomial	
	(1)	(2)	(3)	(4)	(5)
<b>Placebo 1950/51</b>					
<b>Without Covariates</b>					
Treatment Variable	0.0056 (0.0149)	0.0098 (0.0158)	-0.0031 (0.0235)	0.0017 (0.0102)	0.0055 (0.0153)
<b>With Covariates</b>					
Treatment Variable	0.0102 (0.0155)	0.0138 (0.0166)	0.0003 (0.0249)	0.0063 (0.0100)	0.0069 (0.0151)
Running Variable	Linear	Quadratic	Cubic	Linear	Quadratic
<b>Placebo 1952/53</b>					
<b>Without Covariates</b>					
Treatment Variable	-0.0060 (0.0088)	-0.0067 (0.0081)	-0.0072 (0.0119)	0.0001 (0.0101)	0.0004 (0.0128)
<b>With Covariates</b>					
Treatment Variable	-0.0041 (0.0083)	-0.0048 (0.0082)	-0.0084 (0.0109)	-0.0001 (0.0088)	0.0039 (0.0112)
Running Variable	Linear	Quadratic	Cubic	Linear	Quadratic

*Notes:* Standard Errors in parentheses. Sample includes twelve month before and after reform cutoff. Significance levels: \* p 0.10, \*\* p 0.05, \*\*\* p 0.01; Pre-Policy Mean: 15.41 percent. Source: Microcensus 2005-2012, own calculations.



Table 4: Heterogeneity by educational level

	Global Polynomial			Local Polynomial	
	(1)	(2)	(3)	(4)	(5)
<b>Non-College</b>					
<b>Without Covariates</b>					
Treatment Variable	0.0064 (0.0091)	0.0117 (0.0131)	0.0213 (0.0174)	0.0085 (0.0095)	0.0158 (0.0133)
<b>With Covariates</b>					
Treatment Variable	0.0065 (0.0096)	0.0133 (0.0138)	0.0227 (0.0175)	0.0093 (0.0103)	0.0175 (0.0139)
Running Variable	Linear	Quadratic	Cubic	Linear	Quadratic
<b>College</b>					
<b>Without Covariates</b>					
Treatment Variable	0.0557 (0.0346)	0.1285*** (0.0373)	0.1505*** (0.0486)	0.0844** (0.0354)	0.1383*** (0.0392)
<b>With Covariates</b>					
Treatment Variable	0.0573* (0.0328)	0.1291*** (0.0350)	0.1409*** (0.0458)	0.0851*** (0.0330)	0.1325*** (0.0375)
Running Variable	Linear	Quadratic	Cubic	Linear	Quadratic

*Notes:* Standard Errors in parentheses. Sample includes twelve month before and after reform cutoff. Significance levels: \* p 0.10, \*\* p 0.05, \*\*\* p 0.01; Pre-Policy Mean: 15.41 percent.

Source: Microcensus 2005-2012, own calculations.

Table 5: Average Treatment effect on the treated (ATT)

	All (1)	College (2)	Non-College (3)
<b>Global Polynomial</b>			
<b>Linear</b>			
ATT	0.0197	0.0682	0.090
Relative Size ATT (in%)	12.82	21.52	7.16
<b>Quadratic</b>			
ATT	0.047	0.1536	0.0185
Relative Size ATT (in%)	30.59	48.48	15.79
<b>Cubic</b>			
ATT	0.0578	0.1677	0.0315
Relative Size ATT (in%)	37.50	52.91	26.95
<b>Local Polynomial</b>			
<b>Linear</b>			
ATT	0.0332	0.1013	0.0129
Relative Size ATT (in%)	21.53	31.19	11.03
<b>Quadratic</b>			
ATT	0.0543	0.1577	0.0242
Relative Size ATT (in%)	35.30	49.76	20.66
Observations	28,531	5,469	22,694
Eligibility (in%)	76	84	72

*Notes:* The ATT is derived by weighting the ITT effects presented in Tables 1 and 4 using the specification including covariates with the share of employed women eligible for the pension or women. The eligibility rates are calculate from the SOEP data. For the ITT estimates we choose the specification equivalent to column 2 in Table 2, with a global quadratic running variable and with control variables. The sample includes twelve month pre and post reform. Participation rates in training before the reform are 0.154, 0.317, 0.117 for all employed women, employed women with college and employed women without college, respectively.

# Appendix Figures and Tables

Table A.1: Regression Discontinuity: Bandwith 6 months

	Global Polynomial			Local Polynomial	
	(1)	(2)	(3)	(4)	(5)
<b>Without Covariates</b>					
Treatment Variable	0.0329** (0.0131)	0.0273 (0.0142)	0.0945*** (0.0204)	0.0307** (0.0130)	0.0559*** (0.0168)
<b>With Covariates</b>					
Treatment Variable	0.0349*** (0.0108)	0.0243 (0.0137)	0.0679*** (0.0134)	0.0305*** (0.0134)	0.0428*** (0.0118)
Running Variable	Linear	Quadratic	Cubic	Linear	Quadratic

*Notes:* Standard Errors in parentheses; clustered at birth month level. Sample includes six month before and after reform cutoff. Significance levels: \* p 0.10, \*\* p 0.05, \*\*\* p 0.01; Pre-Policy Mean: 15.41 percent. Source: Microcensus 2005-2012, own calculations.

Table A.2: Local Polynomial Regression for different Bandwidths

	6m	9m	12m	6m	9m	12m
<b>Without Covariates</b>						
Treatment Variable	0.0307** (0.0152)	0.0307** (0.0130)	0.0235** (0.0120)	0.0559*** (0.0168)	0.0403** (0.0145)	0.0418*** (0.0131)
<b>With Covariates</b>						
Treatment Variable	0.0305*** (0.0108)	0.0321*** (0.0094)	0.0252*** (0.0087)	0.0428*** (0.0118)	0.0367*** (0.0106)	0.0413*** (0.0100)
Running Variable	Linear	Linear	Linear	Quadratic	Quadratic	Quadratic

*Notes:* Standard Errors in parentheses. Significance levels: \* p 0.10, \*\* p 0.05, \*\*\* p 0.01;  
Source: Microcensus 2005-2012, own calculations.

Table A.3: Donut Regression: Leaving out one or two month on each side of cutoff

	Global Polynomial			Local Polynomial	
	(1)	(2)	(3)	(4)	(5)
<b>One month donut</b>					
<b>Without Covariates</b>					
Treatment Variable	0.0126* (0.0067)	0.0231** (0.0102)	0.0214 (0.0137)	0.0149** (0.0064)	0.0241* (0.0130)
<b>With Covariates</b>					
Treatment Variable	0.0142** (0.0061)	0.0307*** (0.0076)	0.0319** (0.0109)	0.0200*** (0.0051)	0.0352*** (0.0103)
Running Variable	Linear	Quadratic	Cubic	Linear	Quadratic
<b>Two month donut</b>					
<b>Without Covariates</b>					
Treatment Variable	0.0172** (0.0080)	0.0307*** (0.0052)	0.0257*** (0.0071)	0.0197*** (0.0071)	0.0407*** (0.0089)
<b>With Covariates</b>					
Treatment Variable	0.0147* (0.0078)	0.0311*** (0.0044)	0.0277*** (0.0044)	0.0191*** (0.0072)	0.0424*** (0.00652)
Running Variable	Linear	Quadratic	Cubic	Linear	Quadratic

*Notes:* Standard Errors in parentheses. Sample includes twelve month before and after reform cutoff. Significance levels: \* p 0.10, \*\* p 0.05, \*\*\* p 0.01.

Source: Microcensus 2005-2012, own calculations.

## Appendix: Theoretical Model

Based on the the theoretical model outlined in section 2 we obtain the following maximization problem:

$$U_G = U(C_{1i}) + \beta U(C_{2i})R_i + \beta^2 U(C_{3i})(T_i - R_i)$$

max  $U_g$  subject to the intertemporal budget constraint:

$$y_{1i} + \beta y_{2i}R_i + \beta^2 y_{3i}(T_i - R_i) \geq C_{1i} + \beta C_{2i}R_i + \beta^2 C_{3i}(T_i - R_i)$$

The Lagrangian is given by:

$$L = U_G + \lambda [y_{1i} + \beta y_{2i}R_i + \beta^2 y_{3i}(T_i - R_i) - (C_{1i} + \beta C_{2i}R_i + \beta^2 C_{3i}(T_i - R_i))]$$

The set of First Order Conditions i :

$$\frac{\partial L}{\partial C_{1i}} = 0 \Rightarrow U'(C_{1i}) - \lambda = 0 \quad (\text{A.1})$$

$$\frac{\partial L}{\partial C_{2i}} = 0 \Rightarrow \beta R(U'_i(C_{2i}) - \lambda) = 0 \quad (\text{A.2})$$

$$\frac{\partial L}{\partial C_{3i}} = 0 \Rightarrow \beta(T - R)\beta^2[\lambda - U(C_{3i})] = 0 \quad (\text{A.3})$$

$$\frac{\partial L}{\partial I_i} = 0 \Rightarrow \lambda[-w_1(S_i) + (R_i\beta + (R_i - T_i)\alpha\beta^2)w_{2I}(S_i, I_i)] = 0 \quad (\text{A.4})$$

Based on the set of First Order Conditions we derive our results. Firstly, comparative statics yield Equation 5 in the main paper, i.e.

$$\frac{\partial I_i}{\partial R_i} = \frac{[\alpha\beta - 1] w_{2I}(S_i, I_i)}{[R_i + (T_i - R_i)\alpha\beta] w_{2II}(S_i, I_i)}$$

The reform effect is positive as long as  $w_{2I} > 0$  and  $w_{2II} < 0$ .  $\alpha\beta - 1$  is negative since  $\beta < 1$  and  $\alpha < 1$ .

Secondly, Equation 6 in the main paper can be derived as illustrated below. Solving Equation A.10, i.e.  $L_I$ ; for  $\beta$ , gives

$$\beta = \frac{1}{R} \left( \frac{w_1(S_i)}{w_{2I}(S_i, I_i)} + \alpha\beta^2(R - T) \right)$$

Taking  $L_{IS}$  and substituting in  $L_I$ :

$$\text{sign } L_{IS} = \text{sign} \left[ \frac{w_{2SI}(S_i, I_i)}{w_{2I}(S_i, I_i)} - \frac{w_{1I}(S_i)}{w_1(S_i)} \right]$$

The above can then be simplified, written in elasticity form and solved for  $\frac{\partial I_i}{\partial S_i}$ . This then ultimately results in Equation 6 in the main paper :

$$\frac{\partial I_i}{\partial S_i} \gtrless 0 \iff w_{2IS}(S_i, I_i) \frac{S_i}{w_{2I}(S_i, I_i)} =: \eta_{w_{2S}, S} \gtrless \eta_{w_1, S} := w_{1S} \frac{S_i}{w_1(S_i)}$$

Lastly, we apply the multivariate chain rule, given that the reform effect is a function of  $R_i$ ,  $S_i$  and  $I_i$ , which in itself is a function of  $s_i$  and  $R_i$ . Consequently, the size of the reform effect with respect to schooling can be derived by splitting up the effect into two components:

$$\frac{\partial}{\partial S_i} \left[ \frac{\partial I_i}{\partial R_i} \right] = \frac{\partial I^2}{\partial R_i \partial S_i} + \frac{\partial I^2}{\partial R_i \partial I_i} \frac{\partial I}{\partial S_i}$$

Using comparative statics, it is possible to firstly derive:

$$\frac{\partial I^2}{\partial R_i \partial S_i} = \frac{\partial I_i}{\partial R_i} \left[ \frac{w_{2SI}(S_i, I_i)}{w_{2I}(S_i, I_i)} - \frac{w_{2SII}(S_i, I_i)}{w_{2II}(S_i, I_i)} \right]$$

The above expression is positive as long as  $w_{2SII}(S_i, I_i)$  is positive, meaning that the larger  $S$ , the slower is the decline in  $w_{2I}(S_i, I_i)$  with  $I$ .

Secondly, it is possible to derive:

$$\frac{\partial I^2}{\partial R_i \partial I_i} = \frac{\partial I_i}{\partial R_i} \left[ \frac{w_{2II}(S_i, I_i)}{w_{2I}(S_i, I_i)} - \frac{w_{2III}(S_i, I_i)}{w_{2II}(S_i, I_i)} \right]$$

With  $w_{2III}(S_i, I_i) \leq 0$  or sufficiently small the above expression will be negative. Considering, the expression derived for  $\frac{\partial I_i}{\partial S_i}$ , we can write:

$$\frac{\partial I^2}{\partial R_i \partial I_i} \frac{\partial I_i}{\partial S_i} \geq 0 \iff \left[ \frac{w_{1I}(S_i)}{w_1(S_i)} - \frac{w_{2SI}(S_i, I_i)}{w_{2I}(S_i, I_i)} \right]$$

Thus assuming that  $\frac{\partial I_i}{\partial S_i}$  is positive, the reform effect consists of a positive direct and negative indirect effect and the overall effect depends on the relative size of the respective effects, as formulated in Proposition III in the main paper.

## Appendix: Theoretical Model with Utility Cost

In the following we relax the key assumptions of the model presented in the main text (Section 2) and show that the propositions hold without the assumption of dynamic complementarities between initial schooling and further human capital accumulation when we introduce utility costs of training as in Blundell et al. (2019).

### 1. Period

Income in the first period is given by:

$$Y_{1i} = w_1(S_i)$$

### 2. Period

Income in period two is given by:

$$Y_{2i} = w_2(I_i)R_i \tag{A.5}$$

### 3. Period

Income in period three is given by:

$$Y_3 = \alpha w_2(I_i)(T_i - R_i) \tag{A.6}$$

Utility over all three periods is given by:



$$U_G = U_{1i} + \beta U(C_{2i})R_i + \beta^2 U_{3i}(T_i - R_i)$$

Where:

$$U_{1i} = U(C_{1i})$$

$$U_{2i} = U(C_{2i}) - a(S_i)I_i$$

$$U_{3i} = U(C_{3i})$$

Based on the empirical estimates of the structural model in Blundell et al. (2019) we assume that individuals face utility cost of training in period 2,  $a(S_i)$  which fall with schooling  $S_i$ , i.e.  $a(S_i) < 0$ .

Individuals maximize  $U_g$  subject to the intertemporal budget constraint which is given by:

$$y_{1i} + \beta y_{2i}R_i + \beta^2 y_{3i}(T_i - R_i) \geq C_{1i} + \beta C_{2i}R_i + \beta^2 C_{3i}(T_i - R_i)$$

The Lagrangian hence is:

$$L = U_G + \lambda [y_{1i} + \beta y_{2i}R_i + \beta^2 y_{3i}(T_i - R_i) - (C_{1i} + \beta C_{2i}R_i + \beta^2 C_{3i}(T_i - R_i))]$$

and the set of First Order Conditions are:

$$\frac{\partial L}{\partial C_{1i}} = 0 \Rightarrow U'(C_{1i} - a(S_i)I_i) - \lambda = 0 \quad (\text{A.7})$$

$$\frac{\partial L}{\partial C_{2i}} = 0 \Rightarrow \beta R(U'_i(C_{2i}) - \lambda) = 0 \quad (\text{A.8})$$

$$\frac{\partial L}{\partial C_{3i}} = 0 \Rightarrow \beta(T - R)\beta^2[\lambda - U(C_{3i})] = 0 \quad (\text{A.9})$$

$$\frac{\partial L}{\partial I_i} = 0 \Rightarrow \lambda[-a_1(S_i) + (R_i\beta + (R_i - T_i)\alpha\beta^2)w_{2I}(I_i)] = 0. \quad (\text{A.10})$$

Based on the set of First Order Conditions we derive our results. Firstly, comparative statics yield Equation A.11 which is equal to Proposition I in the main paper.

$$\frac{\partial I_i}{\partial R_i} = \frac{[\alpha\beta - 1] w'_2(I_i)}{[R_i + (T_i - R_i)\alpha\beta] w''_2(I_i)} \quad (\text{A.11})$$

As before it holds that the reform effect is positive as long as  $w'_2 > 0$  and  $w''_2 < 0$ .  $\alpha\beta - 1$  is negative since  $\beta < 1$  and  $\alpha < 1$ .

Secondly, the effect of training participation with respect to schooling is captured by Equation A.12:

$$\frac{\partial I_i}{\partial S_i} = \frac{a'(S_i)}{\beta[R_i + (T_i - R_i)\alpha\beta] w''_2(I_i)} \quad (\text{A.12})$$

As long as  $a'(S_i) < 0$  the effect will be positive.

Lastly, we can derive the effect of schooling on the reform effect. Equation A.13 shows that the effect is positive as long as  $w'''_2(I_i)$  is positive and  $w'_2(I_i)w'''_2(I_i) > w''_2(I_i)^2$ .

$$\frac{\partial}{\partial S_i} \left[ \frac{\partial I_i}{\partial R_i} \right] = \frac{[\alpha\beta - 1] a'(S_i) [w''_2(I_i)^2 - w'_2(I_i)w'''_2(I_i)]}{\beta[R_i + (T_i - R_i)\alpha\beta]^2 w''_2(I_i)^2} \quad (\text{A.13})$$